Convergent validity of three measures of sexual sadism: Value of a dimensional measure

Nicholas Longpré

University of Montreal

Jean Proulx

University of Montreal

Sébastien Brouillette-Alarie

University of Montreal

Author Note

Nicholas Longpré, University of Montreal, School of Criminology, International Centre for Comparative Criminology, Philippe-Pinel Institute.

Jean Proulx, University of Montreal, School of Criminology, International Centre for Comparative Criminology, Philippe-Pinel Institute.

Sébastien Brouillette-Alarie, University of Montreal, School of Criminology, International Centre for Comparative Criminology, Philippe-Pinel Institute.

Correspondence concerning this article should be addressed to Nicholas Longpré, University of Montreal, School of Criminology, PO Box 6128, Downtown Station, Montreal, Quebec, Canada, H3C 3J7. Contact: nicholas.longpre@umontreal.ca

Abstract

Sexual sadism can be described as the sexual pleasure produced by acts of cruelty and bodily punishment. The most common method for evaluating sexual sadism is clinical evaluation, that is, evaluation based on the diagnostic criteria of nosological instruments such as the DSM or the ICD. It is also possible to evaluate sadistic sexual preferences by phallometry, which provides a physiological measure of sexual excitation by deviant and nondeviant scenarios. The most recently developed evaluation method is the Severe Sexual Sadism Scale (SESAS), a dimensional instrument that has been empirically validated. Despite the availability of all these measurement techniques, very little research has been conducted on their degree of convergence. Consequently, the aim of the current study was to assess the relationship between these three measures of sexual sadism. Our analyses were conducted on a sample of rapists (N = 72), assessed in a maximum-security penitentiary. There was no significant relation between PPG scores and other measures of sexual sadism. There was, however, an important correlation

between SESAS scores and DSM-IV sexual sadism diagnosis. Our results are consistent with other phallometric studies, which reported no difference in the penile responses of individuals diagnosed as sadists and those not diagnosed as sadists. Results and implications for future research are discussed.

Keywords: sexual sadism, phallometry, SESAS, sexual preference.

Introduction

Sexual Sadism

Throughout history, there have been accounts of behaviours that we would now classify as sadistic. One of the first reported sadists was Gilles de Rais (1404-1440), who raped, killed, and mutilated between 40 and 140 children (Wolf, 1980). During his trial, de Rais stated that torture and mutilation procured him immeasurable pleasure (Wolf, 1980). Although the term "sadism" came to be applied to such behaviours following the publication of the Marquis de Sade's (1740-1814) writings, which eroticized violence and cruelty (Hucker, 1997), it was not until the late 19th century, in Austrian psychiatrist's Richard von Krafft-Ebing's *Psychopathia Sexualis* (1886/1998), that "sadism" appeared in the medical literature. Krafft-Ebing (1886/1998) defined sexual sadism as the experience of pleasure as a result of cruelty and corporal punishment directed towards humans or animals, or the desire to humiliate, strike, hurt, and even destroy others in order to experience sexual pleasure.

Brittain (1970) was one of the first to draw up a portrait of sadistic sexual murderers ("lust murderers"). Many of his criteria have since been applied to the definition of sexual murderers. According to Brittain, sadistic sexual murderers are shy, possess poor social skills, feel inferior in their romantic relationships, and entertain sexual fantasies marked by violence. Amongst other things, they exhibit obsessional traits and emotional detachment. According to MacCulloch, Snowden, Wood, and Mills (1983), there is a spectrum of sadistic sexual behaviours, with sadistic sexual murder constituting one endpoint.

Diagnostic Criteria of Sexual Sadism

In North America, the most common diagnostic criteria for sadism are those set out in the American Psychiatric Association's (2013) Diagnostic and Statistical Manual of Mental Disorders, namely:

> A) Over a period of at least 6 months, recurrent and intense sexual arousal from the physical or psychological suffering of another person, as manifested by fantasies, urges, or behaviors and B) The individual has acted on these sexual urges with a nonconsenting person, or the sexual urges or fantasies cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. (American Psychiatric Association, 2013; p. 695).

In Europe, the most common diagnostic criteria are those of the World Health Organization's (1992) International Classification of Diseases (ICD-10). These criteria differ slightly from those of the American Psychiatric Association. The ICD-10 defines sadomasochism as the preference for sexual activities that involve bondage, corporal punishment, or humiliation: "If the individuals prefer to be the recipient of such stimulation this

is called masochism; if the provider, sadism" (World Health Organization, 1992; p. 172). Furthermore, as non-criminal sexuality may well include sadistic behaviors, a diagnosis of sadism requires that such behaviors be the preferred or exclusive form of sexual gratification.

There are a number of significant problems with current definitions of sexual sadism (Proulx, Blais, & Beauregard, 2007). As Marshall and Hucker (2006) point out, there is no consensus on the defining features of sexual sadism, the requisite number of diagnostic criteria, and the relevance of individual criteria (e.g. animal cruelty) (Marshall, Kennedy, & Yates, 2002; Marshall, Kennedy, Yates, & Serran, 2002). Furthermore, in practice, few experts rely on all the DSM sub-criteria (Fedoroff, 2008). As mentioned by Marshall and Hucker (2006), "each researcher [referring to the researchers who participated in their study] chose an idiosyncratic list of criteria which typically included some features from both DSM and ICD, but also included other features not mentioned in either of these texts" (Marshall & Hucker, 2006; p.1). There is thus little consistency in the criteria used, or in the identification of criteria considered essential for the evaluation of sexual sadism (Marshall, Kennedy, & Yates, 2002; Marshall, Kennedy, Yates, & Serran, 2002; Proulx et al., 2007). These inconsistencies greatly hinder the content validity, construct validity, and discriminant validity of measures of sexual sadism (for more details, see Marshall & Kennedy, 2003).

This lack of consensus on the diagnostic criteria for sexual sadism is in part responsible for the diagnosis' poor inter-rater agreement and reliability (Marshall, Kennedy, & Yates, 2002). According to Krueger (2010), more than 120 years after Krafft-Ebing's definition of sexual sadism, acceptable levels of inter-rater agreement still elude our grasp. For example, Nitschke, Mokros, Osterheider, and Marshall (2013) summarized seven studies of the reliability of a diagnosis of sexual sadism, and reported kappas ranging from .14 to .93. In summary, the study

of sexual sadism is fraught with a number of weaknesses, most of which are related to the diagnosis' validity and reliability.

Sadism scales

In response to the weaknesses of classical approaches to the evaluation of sexual sadism, Marshall and Hucker (2006) developed the Sexual Sadism Scale (SSS). To this end, they asked professionals to evaluate the relevance of 35 diagnostic criteria used in studies of sadism. The 17 criteria judged most relevant to the evaluation of sadism formed the SSS (Figure 1).

<INSERT FIGURE 1 APPROXIMATELY HERE>

More recently, Mokros and Nitschke (Mokros, Schilling, Eher, & Nitschke, 2012; Nitschke et al., 2013; Nitschke, Osterheider, & Mokros, 2009) studied the psychometric properties of the SSS. Nitschke et al. (2009) found 50 sadistic sexual offenders in a sample of 535 patients of a maximum-security psychiatric institution. They then drew a complementary group of 50 non-sadistic sexual offenders at random from the remaining 485 cases and conducted Mokken scaling analysis of the SSS scores based on the 100 select cases. Mokken scaling analysis, which is based on item response theory, has been recommended for the study of the latent structure of psychopathological variables (Meijer & Baneke, 2004). Nitschke et al. (2009) found that only 10 of the 17 items in the SSS satisfied the Mokken scaling criteria for internal consistency. Furthermore, they proposed the addition of one new item, *insertion of objects into bodily orifices*. Their revised scale—the Severe Sexual Sadism Scale (SSSS)—comprises 11 items (Figure 2), 10 of which are taken from Marshall and Hucker's (2006) original scale. The first five items are considered core items. Finally, most of the 11 SSSS items are behavioural, and there is only one physiological item (*Offender is sexually aroused by sadistic acts*).

<INSERT FIGURE 2 APPROXIMATELY HERE>

The SSSS has good psychometric properties. Nitschke et al. (2009) report that it exhibits good inter-rater agreement ($k = 0.86$), possesses good internal consistency ($\alpha = 0.75$) and discriminant validity, and is more sensitive than either DSM-IV or ICD-10 diagnostic criteria (Nitschke et al., 2009). A study undertaken to replicate these results reported essentially identical psychometric properties (Mokros et al., 2012). More recently, Nitschke and colleagues changed the acronym for the Severe Sexual Sadism Scale from SSSS to SESAS (Nitschke et al., 2013).

To date, few studies apart from Nitschke et al.'s (2009) original study, have reported any participant obtaining a maximum score of 11 on the SESAS—the maximum typically reported is 7-8. This may reflect the fact that the SESAS was developed with a sample of judicialized psychiatric patients (Nitschke et al., 2009), while most subsequent studies (Cumbleton, Maillet, & Looman, 2012; Mokros et al., 2012; Pflugradt & Bradley, 2011) were conducted with samples of correctional populations. It is also possible that the unusually high proportion of sadistic sexual offenders' reports in the original study [which was due to the fact that their pair each sadist found in their sample with a non-sadistic sexual offender] also affected the psychometric properties of the SESAS. In addition, the majority of SESAS items measure severe sadistic behaviors (e.g. mutilation of victims); those measuring less severe behaviors—e.g. being the active partner in consensual erotic asphyxiation—have been withdrawn when it transitioned from the SSSS to the SESAS.

Phallometry

Phallometry is a common method for the evaluation of sexual preferences. This method measures the variation of blood flow in the penis upon the presentation of deviant and non-deviant sexual stimuli (usually in the form of an audiotape). Penile tumescence, which can be measured in circumference or volume, is then used as an indicator of sexual arousal. This

method has several advantages over other methods, including a more objective portrait of sexual preferences than self-report measures, and evaluation of a physiological correlate of sexual excitation, which is impossible with other methods (Renaud, Proulx, Rouleau, Granger, Fedoroff, et al., 2003). However, it is not free of problems. A recurrent criticism is its poor test-retest reliability and problematic discriminant validity (Marshall & Fernandez, 2000). Despite its limitations, it remains the most widely used method for the physiological evaluation of sexual preferences. Moreover, several studies (e.g. Proulx, Côté, & Achille, 1993; Seto, Lalumière, Harris, & Chivers, 2012) have reported that some of its limitations may be overcome through the use of appropriate evaluation protocols.

Empirical findings on phallometry concerning discriminant validity in rapists are heterogeneous. Some researchers have reported the method to be effective (e.g. Abel, Becker, Blanchard, & Djenderedjian, 1978; Lalumière & Quinsey, 1993), whereas others have found its utility to be low (e.g. Eccles, Marshall, & Barbaree, 1994; Seto & Kuban, 1996). In Lalumière and Quinsey's (1994) meta-analysis, phallometry adequately discriminated between rapists and non-rapists in 9 of the 16 studies analyzed; discrimination increased in parallel with the violence of the phallometry scenarios. In summary, these results indicate that phallometry, despite its limitations, effectively discriminates between rapists and non-rapists in the majority of cases.

Gratuitous violence, humiliation, and aggression are central to sadistic sexual preferences, and phallometry should allow evaluation of these preferences (Marshall, Hucker, Nitschke, & Mokros, in press). Although several studies have investigated this subject, their results vary widely. For example, while some studies (Proulx, 2001; Proulx et al., 2007) report sadistic sexual offenders to exhibit a preference for coercive sexual activities, others (Barbaree, Seto, Serin, Amos & Preston, 1994; Fedora, et al., 1992) do not. According to Proulx et al.

(2007), sadistic sexual offenders are excited by consensual phallometry scenarios, they but are even more excited by scenarios involving rape with violence. Proulx (2001) reported that rapists who satisfied the MTC:R3 criteria for sadism reacted differently to scenarios involving humiliation or sexual violence than did other types of rapists. Specifically, sadistic rapists exhibited higher phallometric rape indices (maximum penile response to deviant stimuli divided by the maximum response to nondeviant stimuli) than other types of rapists for scenarios involving rape with humiliation (rape index = 1.46) and rape with aggression (rape index = 1.35). In summary, violent scenarios that include elements of torture and humiliation discriminate not only between sadists and non-sadists (Marshall et al., in press; Proulx, 2001), but also between rapists excited by the sexual component of rape and rapists excited by the violent and coercive component (Proulx et al., 2007).

Aim of the study

Currently, the SESAS is the instrument for the evaluation of sadism that has been the most extensively validated (Mokros et al., 2012; Nitschke et al., 2009, 2013). However, the studies of the SESAS' psychometric properties (Cumbleton et al., 2013; Mokros et al., 2012; Nitschke et al., 2009; Pflugradt & Bradley, 2011; Wilson, Pake, & Duffee, 2011) have focused on internal consistency, discriminant validity, and inter-rater agreement, and have reported contradictory results. More recent work (e.g. Mokros, Schilling, Weiss, Nitschke, & Eher, 2014) has focused on the SESAS' latent structure, at the expense of the convergent or criterion validity of the instrument. Consequently, the aim of this study was to evaluate the convergent validity of the SESAS. To this end, the SESAS scores of a sample of incarcerated sexual aggressors were correlated to the results of phallometric evaluation and with diagnoses of sexual sadism based on

DSM-IV (American Psychiatric Association, 1994) criteria. These two evaluation methods target the physiological and behavioural correlates of sexual sadism, respectively.

Method

Participants

The participants in this study were 72 sexual aggressors who had offended against adult women (i.e. females of at least 16 years old) and were incarcerated in a Correctional Service Canada facility. To be included in the study, all participants were required to have committed a sexual offence and have a valid phallometric profile. The sociodemographic characteristics of the sample were similar to those of other studies of incarcerated Canadian sexual offenders (e.g. Barbaree, et al., 1994). The majority of the sample was classified as Caucasian (86.1%), with 9.7% classified as Afro-American and 4.2% classified as "other". The majority of the participants were Francophones (90.5%), and were single (58.3%) and unemployed (69.4%) at the time of incarceration. The mean age of the participants at the beginning of their sentence was 31.8 years (*SD* = 8.2). On average, the participants had committed 19.05 (*SD* = 21.26) previous crimes, 3.06 (*SD* = 3.21) of which had been sexual. At the time of their evaluation at the Regional Reception Centre (see below), 5 participants were diagnosed as sexual sadists and 15 were considered to exhibit traits of sexual sadism, using the DSM-IV criteria. Finally, no participant had any intimate relationship with their victim.

Procedure

Data was collected at the Regional Reception Center (RRC), a maximum-security prison in Sainte-Anne-des-Plaines (Quebec) at which all federal inmates serving a sentence in Quebec undergo an initial 4-6 week evaluation, intended to determine the inmate's security risk and treatment needs. All participants in this study signed a consent form that stipulated that the

information collected would only be used for research purposes. Each participant underwent a battery of psychometric tests and semi-structured interviews. The data collected in this way was complemented by data from official sources (police files, victim statements). If the official data conflicted with the information provided by the participant, the former was considered authoritative.

Computerized Coding Sheet for Sexual Offenders

The data for this study was collected using the QIDS (*Questionnaire informatisé pour les délinquants sexuels*), a computerized coding sheet for sexual offenders developed by St-Yves, Proulx, and McKibben (1994). The QIDS has 18 sections and collects the following information: personal information; criminological information related to case management by Correctional Service of Canada; criminal record (juvenile and adult); information on the index offense; information on precrime, crime, and postcrime factors; attitudes towards the victim of the index offense; information related to arrests; victimology; police and forensic information; offender's personal history and psychosexual development; results of psychological and phallometric testing; suicide attempts; DSM-IV diagnoses.

Coding of the Severe Sexual Sadism Scale (SESAS)

As noted above, the SESAS is a validated instrument that evaluates sexual sadism (Figure 2) through 11 items that describe sadistic fantasies and behaviours. In this study, each item was coded as absent (0) or present (1), using the criteria set out in the Manual for the Assessment of Sexual Sadism (Nitschke, Schilling, Eher, & Mokros, 2012). The manual provides detailed descriptions of each item and leaves little leeway for ambiguity or arbitrariness.

The SESAS was coded a posteriori with existing QIDS variables. The coding process was based on consensus ratings. Most SESAS items had direct QIDS equivalents (e.g. *The offender*

humiliated or degraded his victim), but some items were coded on the basis of the responses to 2-3 QIDS questions. For example, SESAS Item 11, *Insertion of objects into body orifices*, was coded positively if the participant had provided a positive response to the QIDS question on vaginal/anal penetration with an object during the crime or to the question on leaving objects in the victim's body cavities.

Phallometric evaluation

Phallometric evaluation consists of measuring penile responses to deviant (adult nonconsensual sex) and nondeviant (adult consensual sex) auditory sexual stimuli (Michaud & Proulx, 2009). The penile responses are measured using a mercury-in-rubber strain gauge, which responds to fluctuations in the circumference of the penis (Marshall & Fernandez, 2000).

Initially, participants watched a pornographic video of approximately two minutes' duration depicting consensual sex between adults. This video was not part of the phallometric evaluation per se, but allowed participants to become comfortable with the procedure and to immerse themselves in the intrusive evaluation process. Following this video, nine audio recordings were presented to the 72 participants. The scenarios described in the recordings were developed by Abel et al. (1978) and translated into French and validated by Earls and Proulx (1986) and Proulx, Aubut, McKibben, and Côté (1994). The scenarios were used in subsequent studies (e.g. Proulx, 2001; Proulx et al., 2007) and produced results coherent with previous research. A recent book chapter reviewing the assessment of sexual sadism stated that these scenarios are "what appears to be the only phallometric test that actually assesses some of the critical features of sadists" (Marshall et al., in press, p. 9).

A female voice is used for the stimuli, and the mean duration of the stimuli is 208.6 seconds (*SD* = 29.8). The nine recordings were composed of two scenarios depicting consensual

sex with an adult woman, two scenarios depicting the rape with physical violence of an adult woman, two scenarios depicting the rape with humiliation of an adult woman, two scenarios depicting nonsexual physical violence against an adult woman, and a neutral scenario, i.e. a scenario with no depiction of violence or sexuality (for more details, see Castonguay, Proulx, Aubut, McKibben, & Campbell, 1993). Scenarios were presented following the RRC protocol. Since these assessments took place in a correctional setting, and not for research purposes, the presentation order was fixed. This allowed for consistency across assessments, which is essential for correctional and legal procedures.

Compilation of results

The raw data on penile response, which represents sexual excitation in terms of millimetres of penile circumference, was used to calculate deviance indices (maximum penile response to a given deviant stimulus divided by maximum nondeviant response). Deviance indices range from 0 to infinity. The closer the index is to 0, the more an offender prefers nondeviant stimuli to deviant ones. An index of 1 indicates that an offender reacted equally to deviant and non-deviant stimuli. An index greater than 1 indicates that the offender reacted more strongly to deviant stimuli than to nondeviant ones. Indices were calculated for rape with physical violence, rape with humiliation, and nonsexual physical violence. Finally, an overall deviance index (maximum penile response to deviant stimuli divided by the maximum response to nondeviant stimuli) was calculated. Thus, for each participant, four deviance indices were calculated: 1) rape with physical violence; 2) rape with humiliation; 3) nonsexual physical violence; 4) overall deviance.

Diagnosis of sexual sadism

The presence of sexual sadism diagnoses were obtained from offenders' correctional files. These diagnoses were given by psychiatrists using the following DSM-IV criteria: A) Over a period of at least 6 months, recurrent, intense sexually arousing fantasies, sexual urges, or behaviors involving acts (real, not simulated) in which the psychological or physical suffering (including humiliation) of the victim is sexually exciting to the person and B) The fantasies, sexual urges, or behaviors cause clinically significant distress or impairment in social, occupational, or other important areas of functioning (American Psychiatric Association, 1994; p.530). Among the 72 participants, 5 offenders met the diagnosis criteria for sexual sadism and 15 offenders exhibited traits of sexual sadism (i.e. met some of the sadistic sub-criteria, but not enough to establish a clear diagnosis). In order to have a sufficient split between sadists and nonsadists, these two groups were merged in the following analyses.

Results

The participants' mean score on the SESAS was 3.74 (*SD* = 1.57; *Mdn* = 4.00; range 0-7). Fifty-eight percent of participants (*n* = 42) obtained a score of at least 4, Nitschke et al.'s (2009) threshold for sexual sadism.

The results of phallometric testing are presented in Table 1. Of the 72 participants, 22 presented a global deviance index score greater than 1.

<INSERT TABLE 1 APPROXIMATELY HERE>

The results of the analyses for the correlation (Pearson's *r*) between the SESAS and the various rape indices are presented in Table 2. No significant correlation was observed between the SESAS and any phallometric rape index. Furthermore, there was no correlation between the deviance index for nonsexual violence and either the SESAS or the rape indices. As expected, phallometric indices were highly correlated with each other; rape with physical violence was

strongly correlated (r =.90, p <.001) to rape with humiliation, and the overall deviance index was strongly correlated to rape with humiliation (r = .93, p < .001) and rape with physical violence (r =.98, p <.001).

<INSERT TABLE 2 APPROXIMATELY HERE>

Further analyses were conducted to confirm the absence of significant statistical relationships between phallometric rape indices and SESAS scores. First, since the sample exhibited a slightly positive asymmetric distribution for both SESAS scores and rape indices, non-parametric analyses (Spearman's rho) were conducted. Second, different types of deviance indices were tested, namely z-score difference indices (Harris, Rice, Quinsey, Chaplin, & Earls, 1992) and relative arousal indices (Laws, 1989). In all analyses, the results were unequivocal; there were no significant statistical relationship between phallometric rape indices and SESAS scores.

Of the 20 participants either diagnosed with sexual sadism or exhibiting sexual sadism traits, 9 obtained a score of at least 4 on the SESAS. The correlation between the SESAS classification and a DSM-IV diagnosis of sexual sadism was moderate to strong (Cohen, 1988) (r_{pb} = .46; p < .001). As pointed out by Rice and Harris (2005), point-biserial correlation coefficients (r_{pb}) tend to be smaller than product-moment correlation coefficients (r). Therefore, the magnitude of our correlation was potentially underestimated.

The results of the Student's t-test on the differences between the mean phallometric scores of the sadists (either diagnosed or exhibiting traits) and non-sadists are presented in Table 3. Although slight differences were observed between the two groups, none were significant. However, our sample was likely too small to detect a significant relationship and be guarded

against chance variation. Therefore, this absence of significant relationship must be interpreted with caution.

<INSERT TABLE 3 APPROXIMATELY HERE>

Finally, the number of analyses performed (15 statistical tests in total) increased our chances of committing type I errors (i.e. incorrect rejection of a true null hypothesis or a false positive). Therefore, we applied the Holm-Bonferroni correction (Wright, 1992). Controlling for multiple testing did not change the *a priori* observed relationships; what had been significant stayed significant and what had been non-significant stayed non-significant.

Discussion

This study was undertaken to analyse the convergent validity of three measures of sexual sadism: phallometry, SESAS, and DSM-IV. The primary finding of the study was the absence of any significant association between phallometric rape deviance indices and either SESAS scores or DSM-diagnoses of sexual sadism. We suggest two explanations for the absence of correlation.

The first explanation suggests that while the SESAS seems more suitable for the measurement of higher intensity sadism, phallometric rape scenarios seem better suited to lower intensity sadism. Each of these procedures could therefore measure a narrow range of sadistic behaviors and fantasies that would fail to empirically converge, even if they are theoretically part of the same underlying continuum. In fact, with the exception of Nitschke et al.'s (2009) initial study, a maximum SESAS score of 11 has rarely been reported (e.g. Cumbleton et al., 2012). The current study is no exception to this rule. It is possible that the high scores infrequently reported reflect unique characteristics of the samples used to develop the SESAS. While that sample was composed of highly sadistic and judicialized psychiatric patients, most subsequent studies have used correctional samples having lower levels of sexual sadism (e.g. Pflugradt &

Bradley, 2011). As it has been previously mentioned, the SESAS primarily targets severe sadistic behaviours (e.g. mutilation), as the items targeting less intense sadism (e.g. consensual active erotic asphyxiation) were withdrawn from the SSS upon its transition to the SESAS. Therefore, current sadism scales may ignore an important part of the coercion spectrum (Knight, 2014). Recently, Knight and colleagues (Knight, 2010; 2014; Knight Sims-Knight & Guay, 2013; Sims-Knight & Guay, 2011) proposed the existence of an agonistic continuum ranging from nonsadistic sexual coercion—what is termed Paraphilic Coercive Disorder (PCD)—to severe sexual sadism. The term "agonistic" captures the ideas of struggle, anguish, and agony present in both PCD and sexual sadism (Knight et al., 2013). While instruments adequately cover the higher part of the spectrum (ranging from bondage and humiliation to serious harm and torture), they do not cover the lower part of the spectrum (coercive fantasies and scaring behaviors). Recently, Longpré, Guay, and Knight (2014), using the MTC Sadism Scale (MTCSS), a scale measuring sadistic behaviors similar to those measured by the SESAS, also encountered difficulty assessing low-intensity sadism. In summary, the SESAS may be useful for the evaluation of severe sadism, but less effective at evaluating lower-intensity sadism.

In contrast, the phallometric stimuli used in this study seemed better suited for the evaluation of lower-intensity coercive (sadistic) behaviours. Scenarios of rape with physical violence contained violence such as slapping and hitting, and scenarios of rape with humiliation contained degrading elements such as insulting or denigrating the victim. These behaviours represent only a part of the violence spectrum, which also extends to more severe sadistic behaviours such as burning, mutilation and torture, which are better captured by the SESAS. Studies that report that the use of phallometric scenarios depicting a wide spectrum of violence

allows for a better discrimination of sadists and nonsadists would support such an interpretation (Harris, Lalumière, Seto, Rice, & Chaplin, 2012; Seto et al., 2012).

The second explanation for the absence of a correlation between SESAS scores and phallometric results is that the fixed environmental settings of the phallometric rape scenarios did not capture the idiosyncratic sexual fantasies of the assessed offenders. Phallometric rape scenarios include precise descriptions of the physical environment and coercive behavior associated with the aggression, leaving very little room for the projection of the offender's specific fantasies.

Paraphilias, especially sexual sadism, are usually ritualized and idiosyncratic (Abel, Blanchard, Barlow, & Mavissakalian, 1975; Berner, Berger, & Hill, 2003; Seto et al., 2012). Although standardized stimuli can measure some components of sadism, it is possible that phallometric evaluation, which uses uniform, predetermined, and standardized audio stimuli, is incapable of capturing individual patterns of sexual excitation (Quinsey, Chaplin, & Varney, 1981). In particular, the stimuli used in this study may only evaluate a limited range of the sexual preferences of sadistic offenders. This would partially explain why the DSM-IV diagnosis of sadism converged with the SESAS diagnosis, which contains less specific depictions of symptoms of sadism, but not with phallometric indices. The use of individualized stimuli would allow phallometry to take into account individual idiosyncratic fantasies—and, by extension, better evaluate sexual preferences—and increase the technique's ecological validity (Renaud et al., 2010). This would, however, come at the cost of decreased standardization.

Finally, as reported above, SESAS scores and DSM-IV diagnoses of sexual sadism were correlated with each other. This is not surprising, considering that both measures are based on similar criteria that tap into components of sexual sadism. However, as discussed above, there

are major concerns with the current diagnosis of sadism (i.e. liberties taken by professionals during assessment due to an absence of clear guidelines, and an un-nuanced approach to severity) that affect both the validity and fidelity of the assessment. In contrast, the SESAS offers a standardized assessment of sadism that also respects its latent structure. Moreover, the SESAS presents psychometric properties that are superior to those offered by the DSM diagnosis. Given the meaningful consequences of a diagnosis of this magnitude, it is important to ascertain that the tools for the assessment of sadism meet rigorous empirical standards that reflect the condition's latent structure. In this respect, the SESAS seems to outperform the current DSM diagnosis.

Conclusion

This study was undertaken to evaluate the convergent validity of three complementary measures of sexual sadism: the SESAS, phallometric evaluation, and the DSM-IV. The results clearly demonstrate an absence of convergence between the SESAS scores and deviance indices obtained from phallometry. Although quite surprising, there are two possible explanations for this finding: 1) the SESAS and phallometry targeted opposite ends of the sexual sadism continuum; 2) phallometry did not assess idiosyncratic sadistic fantasies. Our study suggests that using more violent and coercive idiosyncratic scenarios is a promising approach to the evaluation of the sexual preferences of sadistic sexual offenders. Furthermore, it appears that the SESAS would benefit from the addition of items that capture low-intensity sexual sadism.

Our study did have some limitations. First, phallometric testing of sexual interest in the rape of adult women is known to have reliability issues (see Marshall & Fernandez, 2003 for a discussion of this issue). By definition, measures deemed unreliable cannot exhibit good convergent validity with other standardized measures—in this study, the DSM-IV and SESAS. However, under the right conditions and with the appropriate scenarios, phallometric testing can

provide valid results (Marshall et al., in press; Seto et al., 2012). Second, the sample size of our study was a bit low. With less than 100 participants, the risk of the sample being unrepresentative or the statistical power being problematic is quite real. Further studies should aim for larger sample sizes. Finally, the prevalence of sadistic rapists in our sample is higher than expected (see Proulx et al., 2007 for a discussion of this issue). This artificial inflation of the prevalence of sadistic offenders may be due to the threshold of the SESAS, the a posteriori coding in our study, or the nature of our sample. Further studies should be undertaken to clarify the appropriate threshold for the SESAS.

In conclusion, future research should focus on the convergence between sexual stimuli and the nature of the acts committed during sexual assaults. What light do sexual preferences measured in the laboratory shed on sadistic behaviours committed during a crime? How accurately do SESAS results reflect offenders' actual sexual preferences? Clearly, further research is necessary to understand the link between sexual preferences and behavioural measures.

References

Abel G. G, Blanchard E. B, Barlow D. H., & Mavissakalian M. (1975). Identifying specific erotic deviations by audiotaped descriptions. *Journal of Applied Behavior Analysis*, 8, 247–260.

Abel, G. G., Becker, D. H., Blanchard, E. B., & Djenderedjian, D. A. (1978). Differentiating sexual aggressive with penile measures. *Criminal Justice and Behavior*, 5, 315-332.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC. : author.

Barbaree, H. E., Seto, M. C., Serin, R. C., Amos, N. L., & Preston, D. L. (1994). Comparisons between sexual and non-sexual rapists subtypes: sexual arousal to rape, offense precursors and offense characteristics. *Criminal Justice and Behavior, 21*, 95-114.

Berner, W., Berger, P., & Hill, A. (2003). Sexual sadism. *International Journal of Therapy and Comparative Criminology, 47,* 383-395.

Brittain, R. (1970). The sadistic murderer. *Medicine, Science and the Law, 10,* 198-207.

Castonguay, L. G., Proulx, J., Aubut, J., McKibben, A., & Campbell, M. (1993). Sexual preference assessment of sexual aggressors: Predictors of penile response magnitude. *Archives of Sexual Behavior,* 22, 325-334.

Cumbleton, A., Maillet, G., & Looman, J. (2012). Diagnosis of Sexual Sadism in a sample of rapists. Paper presented at the 31st Annual Conference of the Association for the Treatment of Sexual Abusers, Denver, CO.

Earls, C. M., & Proulx, J. (1986). The differentiation of francophone rapists and nonrapists using penile circumferential measures. *Criminal Justice and Behavior, 13,* 419-429.

Eccles, A., Marshall, W.L., & Barbaree, H.R. (1994). Differentiating rapists and non-offenders using rape index. *Behaviour Research and Therapy, 32,* 539-546.

Fedora, O., Reddon, J. R., Morrison, J. W., Fedora, S. K., Pascoe, H., & Yeudall, L. T. (1992). Sadism and other paraphilias in normal controls and aggressive and nonaggressive sex offenders. *Archives of Sexual Behavior,* 21, 1-15.

Fedoroff, J. P. (2008). Sadism, sadomasochism, sex and violence. *The Canadian Journal of Psychiatry, 53,* 637-646.

Harris, G. T., Lalumière, M. L., Seto, M. C., Rice, M. E., & Chaplin, T. C. (2012). The sexual arousal of rapists to rape stories: The contributions of sex, nonconsent, and injury. *Archives of Sexual Behavior, 41,* 221-229.

Harris, G.T., Rice, M.E., Quinsey, V.L., Chaplin, T.C. & Earls, C.M. (1992). Maximizing the discriminant validity of phallometric assessment. *Psychological Assessment,* 4, 502-511.

Hucker, S. J. (1997). Sexual sadism: Psychopathology and theory. In D. R. Laws, and W. O'Donohue, (Eds.), *Sexual deviance: Theory, assessment and treatment*. New York, NY: Guilford.

Knight, R. A. (2010). Is a diagnostic category for paraphilic coercive disorder defensible? *Archive of Sexual Behavior*, *39*(2), 419-426.

Knight, R.A. (2014). *Sadism : End of an agonistic continuum*. Oral communication presented at the 33rd conference of the Association for the Treatment of Sexual Abusers (ATSA), San Diego, California.

Knight, R. A., Sims-Knight, J., & Guay, J. P. (2013). Is a separate diagnostic category defensible for paraphilic coercion? *Journal of Criminal Justice*, *41*, 90-99.

Krafft-Ebing, R. von (1998). *Psychopathia sexualis*. Philadelphia: Davis (original work published in 1886).

Krueger, R. B. (2010). The DSM diagnostic criteria for sexual sadism. *Archive of sexual behavior*, *39*, 325-345.

Lalumière, M. L. et Quinsey, V. L. (1993). The sensitivity of phallometric measures with rapists. *Annals of Sex Research*, 6, 123–138.

Lalumière, M. L. et Quinsey, V. L. (1994). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Criminal Justice and Behavior*, 21, 150–175.

Laws, D.R. (1989). Direct monitoring by penile plethysmography. In D.R. Laws (Ed.), *Relapse prevention in sex offenders*. New York : Guilford.

Longpré, N., Guay, J. P. & Knight, R. A (2014). *Investigating the Latent Structure of Sexual Sadism with the MTC Sadism Scale: A Two-parameter IRT Analysis.* Poster presented at

the 33rd conference of the Association for the Treatment of Sexual Abusers (ATSA), San Diego, California.

MacCulloch, M.J., Snowden, P.R., Wood, P.J.W. & Mills, EW., (1983). Sadistic fantasy, sadistic behaviour and offending. *British Journal of Psychiatry*, 143, 20-29. doi:10.1192/bjp.143.1.20

Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review, 20*, 807-822.

Marshall, W. L., & Fernandez, Y. M. (2003). *Phallometric testing with sexual offenders: Theory, research, and practice*. Brandon, VT: Safer Society Press.

Marshall, W. L., & Hucker, S. J. (2006). Issues in the diagnosis of sexual sadism. *Sexual Offender Treatment, 1*, 1-4.

Marshall, W. L., Hucker, S. J., Nitschke, J., & Mokros, A. (under press). Assessment of sexual sadism. In L. A. Craig, & M. Rettenberger (Eds.), *The Wiley-Blackwell handbook on the assessment, treatment, and theories of sexual offending (Volume 2: Assessment of sexual offenders)*. Chichester, UK: Wiley.

Marshall, W. L., & Kennedy, P. (2003). Sexual sadism in sexual offenders: an elusive diagnosis. *Aggression and Violent Behavior, 8*, 1-22.

Marshall, W. L., Kennedy, P., & Yates, P. (2002). Issues concerning the reliability and validity of the diagnosis of sexual sadism applied in prison settings. *Sexual Abuse: A journal of Research and Treatment, 14*, 301-311.

Marshall, W. L., Kennedy, P., Yates, P., & Serran G. (2002). Diagnosing sexual sadism in sexual offenders: reliability across diagnosticians. *International Journal of Offender Therapy and Comparative Criminology, 46*, 668-676. .

Meijer, R. R., & Baneke, J. J. (2004). Analyzing psychopathology items: A case for nonparametric item response theory modeling. *Psychological Methods, 9*, 354-368.

Michaud, P. & Proulx, J. (2009). Penile-response profiles of sexual aggressors during phallometric testing. *Sexual Abuse: A Journal of Research and Treatment, 21*, 308-334. doi:10.1177/1079063209342073

Mokros, A., Schilling, F., Eher, R., & Nitschke, J. (2012). The severe sexual sadism scale: cross-validation and scale properties. *Psychological Assessment*, *24*, 764-769. doi:10.1037/a0026419

Mokros, A., Schilling, F., Weiss, K., Nitschke, J., & Eher, R. (2014). Sadism in sexual offenders: Evidence for dimensionality. *Psychological Assessment*, *26*, 138-147. doi:10.1037/a0034861

Nitschke, J., Ostrheider, M., & Mokros, A. (2009). A cumulative scale of severe sexual sadism. *Sexual Abuse: A journal of Research and Treatment*, *21*, 262-278.

Nitschke, J., Mokros, A., Osterheider, M., & Marshall, W. L. (2013). Sexual sadism: Current diagnostic vagueness and the benefit of behavioral definitions. *International Journal of Offender Therapy and Comparative Criminology*, *57(12)*, 1441-1453. doi:10.1177/0306624X12465923

Nitschke, J., Schilling, F., Eher, R. & Mokros, A. (2012). *SESAS : Manual for the Assessment of Sexual Sadism*. Unpublished document. 1-12.

Pflugradt, D. M., & Bradley, P. A. (2011, November). E*valuating female sex offenders using the cumulative scale of severe sexual sadism*. Poster presented at the 30th Annual Conference of the Association for the Treatment of Sexual Abusers (ATSA), Toronto, Canada.

Proulx, J. (2001). *Sexual preferences and personality disorders of MTC:R3 rapist types*. Oral communication presented at the 19th conference of the Association for the Treatment of Sexual Abusers (ATSA), San Antonio, Texas.

Proulx J., Blais, E., & Beauregard E. (2007). Sadistic sexual offenders. In J. Proulx. M. Cusson, E. Beauregard, and A. Nicole, *Sexual murderers: A comparative analysis and new perspective*. Chichester, UK: John Wiley and Sons.

Proulx, J., Aubut, J., McKibben, A., & Côté, M. (1994). Penile responses of rapists and nonrapists to rape stimuli involving physical violence or humiliation. *Archives of Sexual Behavior, 23*, 295-310.

Proulx, J., Côté, G. & Achille, P. A. (1993). Prevention of voluntary control of penile responses in homosexual pedophiles during phallometric testing. *Journal of Sex Research, 30*, 140-147.

Quinsey, V. L., Chaplin, T. C., & Varney, G. W. (1981). A comparison of rapists' and non-sex offenders' sexual preferences for mutually consenting sex, rape, and physical abuse of women. *Behavioral Assessment, 3*, 127–135.

Renaud, P., Proulx, J., Rouleau, J. L., Bouchard, S., Madrigrano, G., Bradford, J., et al. (2005). The recording of observational behaviors in virtual immersion : A new research and clinical tool to address the problem of sexual preferences with paraphiliacs. *Annual Review of CyberTherapy and Telemedicine : A Decade of VR, 3*, 85-92.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior, 29*, 615-620.

Seto, M.C., & Kuban, M. (1996).Criterion-related validity of a phallometric test for paraphilic rape and sadism. *Behaviour Research and Therapy*, 34, 175–183.

Seto, M. C., Lalumière, M. L., Harris, G. T., & Chivers, M. L. (2012). The sexual responses of sexual sadists. *Journal of Abnormal Psychology, 121,* 739-753.

Sims-Knight, J. E., & Guay, J. (2011). *Is PCD a construct distinct from sadism?* Paper presented at the 30th Annual Conference of the Association for the Treatment of Sexual Abusers (ATSA), Toronto, ON.

St-Yves, M., Proulx, J. & McKibben, A. (1994). *Questionnaire informatisé sur les délinquants sexuels*. Unpublished document, Correctional Service of Canada.

Wilson, R. J., Pake, D. R., & Duffee, S. (2011, November). *DSM-5 Pedohebephilia, PCD, and sadism diagnoses: Reliability in Florida*. Presentation at the 30th Annual Conference of the Association for the Treatment of Sexual Abusers (ATSA), Toronto, Canada.

Wolf, L. (1980). Bluebeard: The life and crimes of Gilles de Rais. New York, N.Y.: C.N. Potter.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: clinical descriptions and diagnostic guidelines*. Geneva.

Wright, S.P. (1992). Adjusted p-values for simultaneous inference. *Biometric*, 48, 1005-1013.

Figure 1

Items in the Sexual Sadism Scale (SSS; Marshall & Hucker, 2006)

1. Offender is sexually aroused by sadistic acts
2. Offender exercises power/control/domination over victim
3. Offender humiliates or degrades the victim
4. Offender tortures victim or engages in acts of cruelty on victim
5. Offender mutilates sexual parts of victim's body
6. Offender has history of choking consensual partners during sex
7. Offender engages in gratuitous violence toward victim
8. Offender has history of cruelty to other persons or animals
9. Offender gratuitously wounds victim
10. Offender attempts to, or succeeds in, strangling, choking, or otherwise asphyxiating victim
11. Offender keeps trophies (e.g., hair, underwear, ID) of victim
12. Offender keeps records (other than trophies) of offense
13. Offender carefully pre-plans offense
14. Offender mutilates nonsexual parts of victim's body
15. Offender engages in bondage with consensual partners during sex
16. Victim is abducted or confined
17. Evidence of ritualism in offense

Figure 2
Items in the Severe Sexual Sadism Scale (SESAS; Nitschke, Osterheider & Mokros, 2009)

1. Offender is sexually aroused by sadistic acts
2. Offender exercises power/control/domination over victim
3. Offender humiliates or degrades the victim
4. Offender tortures victim or engages in acts of cruelty on victim
5. Offender mutilates sexual parts of victim's body
6. Offender engages in gratuitous violence or wounding toward victim
7. Offender keeps records (other than trophies) or trophies (e.g., hair, underwear, ID)
8. Offender mutilates nonsexual parts of victim's body
9. Victim is abducted or confined
10. Evidence of ritualism in offense
11. Insertion of object into bodily orifices

Table 1
Descriptive statistics - Phallometric Scores

	N	Mean (SD)	Min.	Max.
Rape with humiliation	72	.78 (.97)	.05	7.00
Rape with physical violence	72	.81 (1.08)	.03	7.00
Nonsexual physical violence	72	.38 (1.07)	.01	9.00
Global deviance index	72	.91 (1.08)	.05	7.00

Table 2

Convergent validity between Severe Sexual Sadism Scale and phallometric scores (Pearson's r)

	1	2	3	4	5
Total SESAS (1)	1	.04	.07	.04	.11
Rape with humiliation (2)		1	.90***	.03	,.93***
Rape with physical violence (3)			1	.07	.98***
Nonsexual physical violence (4)				1	.05
Global deviance index (5)					1

* = $p < .05$ ** = $p < .01$ *** = $p < .001$

Table 3
Convergent validity between DSM-IV-TR sexual sadism diagnosis and Phallometric scores (Student's t test)

	DSM-IV diagnosis				
	Means		*t*	*N*	Cohen's *d*
	No Sexual Sadism (n = 52)	Sexual Sadism (n = 20[+])			
Rape with humiliation	.68	1.02	-1.33 (n.s.)	72	.36
Rape with physical violence	.70	1.11	-1.48 (n.s.)	72	.38
Nonsexual physical violence	.42	.30	.42 (n.s.)	72	.13
Global deviance index	.78	1.27	-1.77 (n.s.)	72	.45

n.s. = non-significant [+] 5 offenders have a diagnosis and 15 exhibit traits of sexual sadism.